IDENTIFICATION OF THE REGIONS OF THE BOVINE GENOME ASSOCIATED WITH GRAY COAT COLOR IN A NELLORE–ANGUS CROSS POPULATION

A Thesis

by

PAUL WESLEY HOLLAND

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Chair of Committee,	David G. Riley
Committee Members,	James O. Sanders
	Clare A. Gill
	Andy. D. Herring
Head of Department,	H. Russell Cross

May 2015

Major Subject: Animal Breeding

ABSTRACT

The genetics of coat color for cattle are important to breeders and breed associations because phenotypes of these animals are used for breed recognition and premiums or discounts can be given due to the phenotypes. The gene for gray coat color has been determined in other species, but not in cattle. Gray in cattle is known to be recessive based upon observed inheritance. The objective of this study was to identify the regions of the bovine genome associated with gray coat color in a population of Nellore-Angus crossbred cattle. Additionally, proportions of each color and spotting were of interest.

Animals (n = 1941) were classified into phenotypic color categories (i.e. red, black, gray, etc.). Proportions of each color group out of the population were determined, and the proportion of those phenotypes that have any form of spotting. Two genome-wide association analyses were conducted, one where phenotypically gray vs. not gray cattle were analyzed and another where cattle that were very light in color but had a reddish tinge were included as gray. Analyses used Bonferroni correction at $\alpha = 0.05$ ($\alpha/n_{tests} = 1.49 \times 10^{-6}$). Analysis of gray vs. not gray yielded one significant SNP marker on BTA6 at a location of 68,059,441 bp ($P_{raw} = 9.69 \times 10^{-7}$, $P_{adjusted} = 0.032$) (UMD_3.1, NCBI project 32899, Gen Bank GCA_000003055.3). For the analysis of gray and reddish tinged vs. not gray, there were 5 significant markers all on BTA6 forming a region from 62.93 Mb to 83.92 Mb (UMD_3.1, NCBI project 32899, Gen Bank GCA_000003055.3).The same SNP marker from the first analysis was present in

the second, but had an increased significance ($P_{raw} = 1.50 \times 10^{-10}$, $P_{adjusted} = 5.02 \times 10^{-6}$). The region on BTA6 ruled out *syntaxin-17* (*STX17*) on BTA8 and *premelanosome protein* (*PMEL*) on BTA5, previous gray candidate genes from other species, but includes genes such as *v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog* (*KIT*), which is known to cause white coloration (spotting), and *platelet-derived growth factor receptor alpha polypeptide* (*PDGFRA*), the strongest candidate gene for the reddening in Nellore-Angus cattle, and *corin serin pepsidase* (*CORIN*), known for lighter coloration.

DEDICATION

This thesis is dedicated to the memory of my father, Mike Holland. I miss him every day. He was the hardest working man I know, who did everything for the sake of his family. Without him I would not have had the opportunities in life that I have had and I thank him for that. He always supported me in my choices and was always asking me about school. He was always on my mind while working on this research and the examples he set for me as a person kept me working and prevented me from giving up.

ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. David Riley, and my committee members, Dr. James Sanders, Dr. Clare Gill, and Dr. Andy Herring, for their patience and guidance throughout this graduate program. They were always willing to help, be it the countless hours of looking at pictures of the cattle and discussing certain colors with them or clearing up any confusions. I was always greeted with a smile and felt welcome when I took advantage of their "open door" policy. I can never repay them enough for everything they have done during my time here.

I would also like to thank my family for the love and support they have given me. My mother, Ginger Holland, and brothers, John, Matt, and Brett Holland, by offering constant support and encouragement, that kept me pushing throughout this tough year. I cannot adequately thank them.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

INTRODUCTION

The genetics of coat color for cattle and other species are important to breeders and breed associations because the phenotypes of these animals are used for breed recognition. Also premiums or discounts due to the coat color of cattle have an impact on the breeding strategies of producers, therefore understanding the genes and their inheritance is desired (Hulsman Hanna et al., 2014).

Coat color determination in other species such as horses, mice, and humans may help determine how color, in this case gray, is inherited in cattle and could also give a homologous location to look for the causative gene in the bovine genome. In all three species, gray is a mutation that is actually a premature graying caused by loss of melanocytes (Martínez-Esparza et al., 1999; Kavar et al., 2012). This could also be the case in cattle as well, as some cattle are born red and start to gray as they age, some faster than others. In the case of mice and humans, graying is caused by the *premelanosome protein* (*PMEL*) gene as described by Martínez-Esparza et al. (1999). If gray coat color in cattle is the result of the same gene, it would give rise to the theory that gray in cattle could be a type of dilution, because in studies by Kühn and Weikard (2007) and Schmutz and Dreger (2013), mutations in the *PMEL* gene caused dilutions of color in Charolais crosses and Highland or Galloway cattle, respectively.

The objective of this study was to identify the regions of the bovine genome associated with gray coat color. Gray coat color is recessive based upon observed inheritance, therefore it is most likely caused by a major gene. Genome-wide association

analyses were conducted using SNP genotypes of a population of Nellore-Angus crossbred cattle. Two analyses were completed, one with phenotypically gray vs. not gray cattle and another which included cattle that are very light in color but had a reddish tinge in the gray category instead of the not gray category.

LITERATURE REVIEW

Basic Inheritance of Coat Color in Cattle

The absence or presence of melanin pigment in the hair or skin of cattle is the basis for coat color. Melanin can be found inside melanosomes of the cytoplasm in the melanocytes of cattle. As described by Wasmeir et al. (2008), melanosomes are large organelles (~500 nm in diameter) that are the cellular site of storage, transport, and synthesis of melanin pigment. Melanosomes are synthesized in skin melanocytes in mammals. There are two types of melanin found in cattle and other mammals: pheomelanin and eumelanin. Eumelanin is responsible for black and brown colors, pheomelanin for reddish brown, reds, tans and yellows, and white hair occurs where there are no melanocytes in the hair or skin (Olson, 1999). In the early stages of embryonic development, the melanocytes begin migration from the neural crest to the rest of the body; the furthest parts of the body, the tip of the tail, nose, and feet, can be white due the melanocytes not reaching those parts of the body. Coat color in cattle is also determined by many different loci. Loci that affect cattle coat color, based upon observed inheritance, are the extension, dilution (*D* in Charolais, *C* in Simmental), roaning (*R*), spotting (*S*), graying (*G*), reddening (*N*), black nose, feet and tail (*B*), and agouti (*A*) loci (J. O. Sanders, Texas A&M University, Dept. of Animal Science, personal communication). Differences in alleles at the extension locus (E^D, E^+, e) result in differentiation of black and red base coat colors. The two types of dilution loci result in lighter shades of the base colors. The presence of the *D* allele in Charolais type cattle

causes a dilution to white in homozygotes and a lighter shade in heterozygotes. The presence of the *c* allele in Simmental type cattle can result in a dilution to a lighter color. Presence of the roaning allele causes white cattle in homozygotes and a roan color in heterozygotes (mixture of pigmented and non-pigmented hairs). In cattle, spotting alleles can cause various types of white spotting (S^H, S^S, S^G, S^P, *s*). Combinations of these alleles cause spotted phenotypes, e.g., "full" Hereford pattern, bald face, speckled, spotted (more or less, random spots), or "full" Simmental pattern. The reddening locus is linked to the spotting locus and causes cattle with a heterozygous black (E^DE^+) genotype to be a color lighter than black and appear to be "reddened" (Hulsman Hanna et al., 2014). Presence or absence of different alleles at the agouti locus determine whether or not the animal is banded (brindle) or a solid color as described by Olson (1999). If the *B* allele is present in red or gray colored cattle, their nose, feet and tail will be black in homozygotes and only some degree of black pigmentation in heterozygotes. Some, but not all, of these genes have been characterized in the bovine genome.

Gray Color Variations in Cattle

Gray or variations of gray coat color can typically be seen in *Bos indicus*, Italian, and Eastern European breeds. Gray color is caused by homozygosity of the *g* allele in cattle that are *ee*, E^+E^+, or E^+e (red base color) genotypes at the extension locus. Also, these cattle commonly carry the allele for the black nose, feet and tail, and most of the homozygotes and heterozygotes will have black hairs intermingled among the gray or red hairs. White-spotting patterns can be seen in many of these breeds as well. The Dhanni of Pakistan and the Landim of Mozambique have a pattern that is predominantly

white but there is small amounts of black spotting and the ears and muzzle are black as well. These cattle are not "gray" but have a grayish coloration. This pattern can also be seen in British White Park cattle and the Blanco Orejinegro of Colombia. According to Olson (1999), these patterns are possibly from a modification of the line-backed pattern. Durkin et al. (2012) showed that lineback or color-sided is due to an insertion of *v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog* (*KIT*) on BTA29. Another type of cattle, the Bapedi, a South African breed, have a mottle gray/black color which is not normally observed in other breeds.

Genomic Regions Known to Influence Cattle Coat Color

Based upon observed inheritance, *melanocortin 1 receptor* (*MC1R*) corresponds to the extension locus; E^D, E^+, and *e* alleles and follow a dominance model of $E^D > E^+ >$ *e*. The E^D allele is caused by a point mutation of leucine to proline and creates a receptor that results in eumelanin production, resulting in black coat color, whereas the *e* allele only produces pheomelanin, because of a non-functional receptor resulting in red coat color. The E^+ allele can produce red coat color, but it produces a receptor that responds to the α-melanocyte stimulating hormone (α-MSH) and agouti-signaling protein (ASP); it leads to full range of color from yellow to black. Klungland et al. (1995) was the first to report a mutation in Norwegian and Icelandic cattle where black cattle (heterozygous E^D at extension) were phenotypically red. Joerg et al. (1996) developed a test for the recessive allele that causes red coat color in black Holsteins by PCR amplification of the bovine *MC1R* gene (originally known as *melanocyte-stimulating hormone receptor, MSHR*). Analysis of the sequences revealed a deletion of

the G-residues in the red Holstein at position 771 or 772. This represents a frameshift mutation that leads to a premature stop codon, which truncates *MC1R* resulting in a nonfunctional receptor. This gene was mapped to bovine chromosome 18 using 36 bovine-hamster somatic cell hybrid panels by Werth et al. (1996).

Variant red color in cattle looks the same as the traditional red coat color phenotype, but instead has dominant inheritance and does not rely on *MC1R*. Dreger and Schmutz (2010) performed a study to determine whether this rare variant red phenotype co-segregates with *MC1R*, *agouti signaling protein* (*ASIP*), *attractin* (*ATRN*), *melatonin receptor 1A* (*MTNR1A*), or *beta-defensin 103* (*DEFB103*). *Melatonin receptor 1A* and *DEFB103* are both located on BTA27 within 10Mb of each other. After analyzing data from both the family that traced back to the original animal with this phenotype and the SNP, it was determined that *DEFB103* was the causative gene for the variant red phenotype.

Olson (1999) discussed three different types of *ASIP* (*agouti signaling protein*) and the effects it has on the coat colors: patterned blackish (A^{bp}), white-bellied (a^{w}), and fawn (a^{i}). Patterned black is similar to the wild-type coloration, but cattle are nearly completely black, white-bellied is the removal of most of the red pigmentation and part of the black hair pigmentation particularly on the sides of cattle, and fawn is described as the removal of both red and black hair pigmentation on the underline and back. Giradot et al. (2006) determined that another type of *ASIP* (a^{br}) was responsible for the brindle phenotypes of Normande, a French breed of cattle. Brindle was described as alternating

stripes of red and black hair pigmentation. *Agouti signaling protein* was assigned to BTA13 through somatic cell hybrid analysis by Schläpfer et al. (2001).

Dilution in Charolais x German Holstein F_2 cattle was mapped to BTA5 close to the *ETH10* marker by Kühn and Weikard (2007). It was concluded that there are 2 interacting mutations that affect coat color dilution in Charolais, because the *e* allele is almost fixed in this breed. There was high diversity of gray looking (smokey) and red/yellow dilute phenotypes observed in the F_2 cattle, which caused these researchers to suspect 2 mutations in *premelanosome protein* (*PMEL*; also known as *PMEL17* or the *silver homolog*, *SILV*) were responsible for dilution in these types of cattle. Schmutz and Dreger (2013), showed that the "dun type" dilution (dilution of black or red base color taao a lighter shade of red) in Highland cattle and Galloway cattle is associated with a deletion of leucine in the signal peptide region of *PMEL*.

Berryere et al. (2003) determined that the *tyrosinase-related protein 1* (*TYRP1*) gene is not an influence in the diluted coat colors in Simmental, Belted Galloway, and Charolais/Angus sired offspring, nor is the gene responsible for the brown coat color seen in Braunvieh cattle. In Dexter cattle, on the other hand, there was an amino acid change in the protein produced by the diluted cattle. This change is consistent with a mutation that replaces histidine with tyrosine, which alters the shade of eumelanin pigment. It was concluded that *TYRP1* mapped to BTA8 based upon the location that *TYRP1* was previously mapped in humans, mice, horses, and the significant LOD-scores for markers on BTA8.

Seitz et al. (1999) determined that the roan phenotype in Belgian Blue and Shorthorn cattle was caused by a missense mutation in the *v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog* (*KIT*) *ligand* (*KITLG*) gene (originally known as *mast cell growth factor*, *MGF*). This mutation replaces a hydrophobic residue with an acidic residue in the 193rd amino acid of the *MGF* locus in cattle. It is unknown whether the mutation is linked to another mutation that could be the actual cause.

Using a high-density (774,660) single nucleotide polymorphism (SNP) chip for a genome-wide association study, Philipp et al. (2011) determined that a missense (R210I) mutation in the highly conserved basic region of *microphthalmia-associated transcription factor* (*MITF*) on bovine chromosome 22 causes the dominant white phenotype in German Fleckvieh cattle.

Reinsch et al. (1999) discovered close synteny of the *KIT* locus with a QTL for the degree of spotting, which identified this locus as the top candidate gene for the proportion of unpigmented coat in cattle for breeds with recessive spotting. *V-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog* was mapped to chromosome 6 in cattle between the ILSTS097 and CSN3 markers (Barendse et al., 1997), and the QTL for degree of spotting was concordant with this location (Reinsch et al., 1999). The spotting locus was found to be located on chromosome 6 in Hereford crossbred cattle, between the BMS2460 and BM4528 microsatellites (Grosz and MacNeil, 1999). Reinsch et al. (1999) hypothesized a guideline for future research; the authors suggested that most likely there was a series of dominant and recessive alleles at the *KIT* locus that

correspond with the already known alleles, based upon observed inheritance, at the spotting locus.

In a study by Drögemüller et al. (2009), significant linkage between the belted locus and BTA3 markers in Brown Swiss was found, which ruled out the previous candidate genes *KIT*, *ADAM metallopeptidase with thrombospondin type 1 motif, 20* (*ADAMTS20*), and *endothelin receptor type B* (*EDNRB*), because they are not located on BTA3. The belted locus was fine-mapped to a 922-kb region by identification of recombinant haplotypes from 2 apparent recombination events that were inherited by offspring without belts from their sires with belts. Some other predominant breeds seen with the belted phenotype are the Dutch Belted and Galloway. Drögemüller et al. (2010), looked at these other belted phenotype breeds and saw there was a high degree of homozygosity in Dutch Belted and Belted Galloway in the interval on BTA3, suggesting that the phenotype is caused by the same locus as in the Brown Swiss. A candidate gene for this belted phenotype is *hes family bHLH transcription factor 6* (*HES6*). This gene is located within the interval on BTA3 and encodes a developmental transcription factor.

In a study by Schmutz et al. (2003), albinism in cattle was determined to be due to a frameshift mutation in *tyrosinase* (*TYR*). *Tyrosinase* production is regulated by the extension locus, which determines whether eumelanin or pheomelanin (black or red) pigment is produced (Olson, 1999). An albino heifer and her phenotypically normal dam were sequenced. The amplification of *TYR* in that heifer and her dam indicated that an

additional cytosine in a run of four cytosines was observed in the albino heifer, and the dam was heterozygous for this variant. This inserted cytosine in the albino calf caused the frameshift mutation that resulted in a stop codon at residue 316, causing deregulation of *TYR*. *Tyrosinase* was also mapped to bovine chromosome 29 by Schmidtz et al (2001).

Table 1 below lists known genes that influence coat color, the chromosome they have been mapped to, along with the starting position in the bovine genome. Starting positions in the bovine genome were obtained from the UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway) using the Nov. 2009 *Bos Taurus* draft assembly (UMD_3.1, NCBI project 32899, Gen Bank GCA_000003055.3).

Table 1. Genes that influence color and respective locations

Study	Gene	Chromosome:Mb	Coloration/pattern
Drögemüller et al., 2009 Drögemüller et al., 2010	*HES6*	BTA3:118.18[6]	Belted pattern
Seitz et al., 1999	*KITLG*	BTA5:18.31[6]	Roan
Kühn and Weikard, 2007 Schmutz and Dreger, 2013	*PMEL*	BTA5:57.67	Dilution[1] dun (dilution)
Martínez-Esparza et al., 1999 Kwon et al., 1991	*PMEL* [2]	BTA5:57.67[6]	Gray[2]
Enshell-Siejffers et al., 2008	*CORIN* [5]	BTA6:67.93[7]	Lighter color[5]
Reinsch et al., 1999	*KIT*	BTA6:71.80	Spotting
Berryere et al., 2003	*TYRP1*	BTA8:31.71	Dun (brown)
Kavar et al., 2012	*STX17* [3]	BTA8:65.43[6]	Gray[3]
Olson, 1999 Girardot et al., 2006 Albrecht et al., 2012	*ASIP*	BTA13:64.21	Patternedblack white-bellied fawn/tan brindle
Norris and Whan, 2008	*ITCH* [4]	BTA13:64.36[6]	Dominant white[4]
Joerg et al., 1996	*MC1R*	BTA18:14.76	Base color (black/red)
Phillip et al., 2011	*MITF*	BTA22:31.74	Dominant white
Dreger et al., 2010	*DEFB103*	BTA27:4.90	Variant red
Schmutz et al., 2003	*TYR*	BTA29:63.52	Albinism

[1]Dilution in Charolais cattle
[2]Gene and color associated with gray in mice and humans
[3]Gene and color associated with gray in horses
[4]Gene and color associated with white in sheep
[5]Gene and color associated with lighter coat color in mice
[6]Homologus bovine chromosome location found from UCSC Genome Browser http://genome.ucsc.edu/ using assembly UMD_3.1
[7]In ENSEMBL Genome Browser annotations as ENSBTAG00000002199

What is Known about Gray Coat Color in Other Species

Similar to cattle, horse color inheritance can be determined by many different loci, such as the extension locus (*E* and *e* alleles), agouti/restriction locus (*A* and *a* alleles), roaning locus (*R* and *r* alleles), graying locus (*G* and *g* alleles), and also two dilution loci, non-lineback dilution (*C* and *c* alleles) and lineback dilution (*D* and *d* alleles). Horses also have loci for dominant white (*W* and *w*), tobiano, (*T* and *t*), frame (*F* and *f*), sabino (*Sb-1* and *sb-1*), splashed white pattern (*Spl* and *spl*), Appaloosa complex (*Ap* and *ap*), and pattern (*Pn* and *pn*) (Sponenberg, 2009). At the extension locus, differences in genotypes result in either bay, black, or sorrel base colors. This locus is also linked to the tobiano, sabino-1 and roaning loci. In horses that are homozygous or heterozygous *E* at the extension locus, bay versus black color is determined by the agouti/restriction alleles. One of the types of dilution in horses results in the dilution of bay to buckskin, sorrel to palomino, and sometimes black to smokey, if the horses are heterozygous for the c allele; the other type of dilution causes dilution of bay to dun, black to grullo, and sorrel to red dun in the presence of the *D* allele.

Graying in horses is a progressive graying of the base color and is caused by the presence of the *G* allele. When foaled, horses can be any color (i.e. sorrel, bay, black), but become gray prematurely due to the loss of hair pigmentation. Foal color does not predict the coat color of the horse, but foals that are born dark black will gray out and foals that are born a dull black will not (J. O. Sanders, Texas A&M University, Dept. of Animal Science, personal communication). The face of the horse grays first and then the rest of the body. This premature graying with age is an autosomal dominant trait that is

due to a 4.6-kb duplication in the intron 6 of the *syntaxin-17 (STX17)* gene (Kavar et al., 2012). This mutation is also associated with high incidences of skin pigmentation loss and melanoma as well. A new method for detecting the gray coat color allele in horses has been developed and verified on subjects from 30 non-gray and 30 gray horses (Kavar et al., 2012). Most of the animals in that study originated from Slovenian horses (*gg*), Lippizan horses (*GG, Gg, gg*), and Thoroughbreds (*GG, gg*). Because the marker was a dominant character, the method only allowed discrimination between gray and non-gray horses, but not between homozygous and heterozygous gray horses (Kavar et al., 2012).

In mice, premature graying is due to the loss of follicular melanocytes. This "silver" phenotype is caused by a mutation generating a truncated transcript of the murine *SILV* gene. Humans have a homologous gene, *PMEL*, and its products were potential markers of human melanoma and immunotherapy targets (Martínez-Esparza et al., 1999). This melanocyte-specific gene, was mapped near the silver coat locus in mice. A clone of the *PMEL* gene, *PMEL17-1*, was used to map the murine gene. A panel of mouse-hamster somatic cell hybrids were analyzed. Two of the hybrids retained mouse chromosome 10 which indicated that the murine *PMEL* was located on chromosome 10. The human gene was mapped in a similar way using hamster-human somatic cell hybrids and it was determined that *PMEL* was on chromosome 12. Human chromosome 12 is known to share conserved synteny with mouse chromosome 10 (Kwon et al., 1991).

Coat color of sheep due to *ASIP* results in three different phenotypes. Dominant white or tan *ASIP* allele (A^{wt}) is responsible for yellow or red phenotypes, non-agouti (A^{a}) results in black or brown phenotypes, and badgerface (A^{b}) characterized by pale dorsal phaeomelanic and a darker ventral eumelanic pattern (Norris and Whan, 2008). To characterize the *ASIP* locus, sequence analysis of genomic DNA, BAC clones, and RT-PCR products from Texel, Merino, Barbary, and Romanov sheep was used. Through RT-PCR analysis of white and recessive black Merino sheep, it was indicated that the dominant white phenotype was caused by high levels of deregulated expression of *agouti* from an *ITCH* (*itchy homolog E3 ubiquitin protein ligase*) gene promoter (Norris and Whan, 2008).

Another gene in mice, *corin serine pepsidase* (*CORIN*), results in a lighter shade of coat color. This light shade phenotype depends on the alleles of *ASIP*. This phenotype will be expressed if a functional allele (*A*) of *ASIP* is present, but *CORIN* mutants that are homozygous for the null allele at *ASIP* (*a*/*a*) are black and not able to be distinguished from the wild type (Enshell-Seijffers et al., 2008). The *CORIN* mutants also had increased basal band lengths when compared to the wild types. This increase basal band length resulted in an increase in the amount of pheomelanin in the hair, which is responsible for the light shades of color (reddish brown, red, tan, yellow). These "tipped" hairs have also been seen in Indian cattle and the cattle exhibiting a solid gray pattern seem to have longer white tips on the hairs (Rhoad, 1936).

Statistical Software

Software considered for analyses were PLINK (Purcell et al., 2007), the Q-K procedures of JMP Genomics (SAS Inst. Inc., Cary, NC), or GEMMA (Zhou and Stephens, 2012). These software can all be used for genome-wide association studies (GWAS).

PLINK was designed to handle large data sets more conveniently and rapidly than other tools used for genetic analysis. PLINK has 5 main functions as described by Purcell et al. (2007): the management of data, provide summary statistics, population stratification, association analyses, and estimate identity-by-descent.

JMP Genomics can be used for analyses ranging from simple case-control associations to linear models that have covariates, interactions, and random effects. In terms of some of its GWAS capabilities, SNP associations and interactions can be examined. The Q-K procedures of this software allows for simplification in the creation and integration of relationship matrices into associations studies.

GEMMA tests for associations in GWAS through the use of linear mixed models. GEMMA implements the Genome-wide Efficient Mixed Model Association algorithm for linear mixed models. For marker association tests with only a single phenotype and for estimating the proportion of variance in the phenotypes explained (PVE), it fits a univariate linear mixed model. If the marker association test has multiple phenotypes at the same time and genetic correlations among complex phenotypes are needed, a multivariate linear mixed model is fitted.

OBJECTIVE

The objective of this study was to:

Identify the regions of the bovine genome associated with gray coat color in Nellore–Angus cross cattle.

MATERIALS AND METHODS

Cattle Populations

All cattle used for the study were produced at the Texas A&M AgriLife Research Center located near McGregor, Texas. The population consisted of Nellore-Angus crossbreds. Cycle 1 calves (n = 480) were part of 14 full-sibling Nellore-Angus embryo-transfer F_2 families. The families were produced from 4 F_1 Nellore-Angus sires and 13 F_1 Nellore-Angus cows. The recipient cows were crossbred *Bos indicus-Bos taurus* cows (n = 229); most of these cows had one calf, but some had up to 4 calves. The 4 Nellore-Angus F_1 sires were additionally mated to an F_1 and F_2 population *Bos indicus* (Brahman) × *Bos taurus* (Hereford or Angus) cows to produce natural service paternal half-sibling families (n = 266 calves). Calves were born in either the spring or fall from spring of 2003 to spring of 2007. Three additional populations were made as crosses using the Nellore and Angus breeds. Another F_2 group (Cycle 2) consisted of all possible reciprocal combinations of Nellore (N) and Angus (A) (NA × NA, AN × AN, NA × AN, and AN × NA), where the sire breed is represented first and dam breed second (NA, Nellore sired and Angus dam, AN, Angus sired and Nellore dam). They were produced by natural service from 2009 to 2014. Bull and cows from the original embryo transfer population were mated to produce F_3 calves (Cycle 3); there were a total of 169 calves produced in Cycle 2 and Cycle 3 with records available, but only some of the steers were genotyped (n = 70). A total of 142 F_4 calves were produced in Cycle 4 in 2014. These calves were made from mating F_3 bulls to F_3 cows .All procedures involving animals

were approved by the Texas A&M University Institutional Animal Care and Use Committee.

Phenotypes

After examination of available photographs and any recorded color for animals, categories of gray phenotypes were created. All cattle from these populations were classified into color groups (i.e. red, black, gray, gray brindle, red brindle, etc.) and the proportion of each phenotype in the population was determined. For the association analyses, these cattle were grouped into two categories, gray (similar to "cases" in disease studies) (n = 42) and not gray (similar to "controls" in disease studies). Another analysis of these data was conducted after classifying those cattle that are very light in color but with a reddish tinge coat color as gray (n = 53). Figure 1 shows typical cattle placed in the two gray categories for analysis.

Figure 1**.** Gray phenotype classification for analysis. A, B, and C are shades of gray that were put into the first analysis of gray. D, E, and F are shades of red tinged gray that were included in the second analysis of gray

Genotypes

Genotypes on these animals were available from previous studies. Blood samples were collected at weaning on the calves for isolation of DNA as described by Riley et al. (2013). For Cycle 1 cattle, 200 mL of blood were collected, and for both Cycle 2 and Cycle 3, 30 mL of blood was collected. Genotypes on these cattle were obtained using the Infinium BovineSNP50v1 assay (Illumina, Inc., San Diego, CA); version 1 chip was used on Cycle 1 blood samples and version 2 chip was used on Cycle 2 and Cycle 3 samples. Version 1 and 2 data were merged using PLINK (Purcell et al., 2007). Usable data for 34,957 SNP loci per animal were produced and on average there were 26,692 informative markers per F_2 family, or 1 informative marker per 101 kb (Riley et al., 2013).

Statistical Analyses

The phenotypes for this study were analyzed using linear mixed models using the univariate procedures of GEMMA (Zhou and Stephens, 2012). GEMMA (Zhou and Stephens, 2012) was chosen over both PLINK (Purcell et al., 2007) and JMP Genomics (SAS Inst., Inc., Cary, NC). GEMMA easily allows the genomic relationship matrix to be incorporated into the model and runs the analyses more rapidly than the others. One association analysis was conducted on cattle that were classified as phenotypically gray vs. not gray and another association analysis was conducted using both the gray colored cattle and the cattle that have a reddish tinged coat color vs. not gray. Association analyses were based on a regression of the gray-not gray variable (0 and 1) on genotypic values corresponding to homozygotes (values of 0 and 2) and heterozygotes (value of 1)

at each SNP locus. *Melanocortin 1 receptor* (*MC1R*) genotypes were fitted as fixed effects in these analyses because cattle that are homozygous or heterozygous E^D at the extension locus (phenotypically black) would not allow expression of gray. The genotypes at *MC1R* were categories, based upon the alleles at *MC1R* (i.e. *ee*, E^+e, E^+E^+, E^De, E^DE^+, E^DE^D). The genomic relationship matrix was also incorporated into the model. Genome-wide association analysis consisted of distinct regressions for each SNP genotype. Using the methods of Bonferroni correction (Bland and Altman, 1995), genome-wide significance level was set to $\alpha = 0.05$ ($\alpha/n_{tests} = 1.49 \times 10^{-6}$).

RESULTS AND DISCUSSION

Phenotypic Proportions

From the available records, either photographs and/or descriptions of the cattle in the populations, the cattle in the population were placed into phenotypic color categories (i.e. red, black, gray, gray brindle, red brindle, etc.). From these categories proportions of each color phenotype were calculated as a fraction of that phenotype out of the total number of animals with records. Also, cattle with any form of spotting, ranging from a few white hairs to paint coloration, were calculated as the fraction of spotted phenotypes out of the number of animals with that color. Animals were excluded from these results if there were no phenotypic data available. Table 2 lists the overall proportions of each color in the population as well as the proportion of spotted animals of each color. Table 3 consists of only the natural service cattle from Cycle 1 (n = 266). Table 4 only consists of the cattle that had been genotyped and were used in the analyses of gray vs. not gray. Approximately half of the population in consideration was phenotypically black; higher proportions of black are typical of these types of crosses. Black is dominant to red at the extension locus meaning that only one allele is needed for black to be expressed, and also black will not allow expression of gray.

Table 2. Phenotypic color and spotting proportions of cattle population[1,2,3]

Color	Count	Proportion[4]	Spotting count	Proportion[5]
Black	833	0.496	237	0.285
Black brindle (br)	60	0.036	19	0.317
Red	68	0.040	17	0.250
Red br	237	0.141	28	0.118
Gray	39	0.023	10	0.256
Gray br	66	0.039	16	0.242
Tan	30	0.018	9	0.300
Tan br	25	0.015	5	0.200
Brown	110	0.065	45	0.409
Brown br	8	0.005	2	0.250
Brownish black	98	0.058	35	0.357
Brownish black br	6	0.004	5	0.833
Reddish black	46	0.027	12	0.261
Reddish black br	21	0.013	4	0.190
Reddish brown	25	0.015	9	0.360
Reddish brown br	8	0.005	1	0.125
Totals	1680		454	

[1]66 animals were excluded for no phenotypic descriptions available

[2]Includes genotyped animals in Table 4

[3]Excludes Cycle 1 natural service cattle in Table 3

[4]These values are a proportion of the total animals

[5]These values are a proportion of the color category

Table 3. Phenotypic color and spotting proportions of Cycle 1 natural service cattle[1]

Color	Count	Proportion[2]	Spotting count	Proportion[3]
Black	127	0.486	35	0.276
Black brindle (br)	7	0.027	2	0.286
Red	9	0.034	3	0.333
Red br	37	0.142	18	0.486
Gray	9	0.034	0	0
Gray br	7	0.027	0	0
Tan	10	0.038	1	0.100
Tan br	7	0.027	1	0.143
Brown	38	0.146	26	0.684
Brown br	2	0.008	1	0.500
Brownish black	1	0.004	1	1.000
Brownish black br	1	0.004	0	0
Reddish black	3	0.011	0	0
Reddish black br	1	0.004	0	0
Reddish brown	1	0.004	1	1.000
Reddish brown br	1	0.004	1	1.000
Totals	261		90	

[1]5 animals were excluded for no phenotypic descriptions available

[2]These values are a proportion of the total animals

[3]These values are a proportion of the color category

Table 4. Phenotypic color and spotting proportions of genotyped animals[1]

Color	Count	Proportion[2]	Spotting count	Proportion[3]
Black	391	0.504	105	0.269
Black brindle (br)	36	0.046	10	0.278
Red	26	0.034	9	0.346
Red br	98	0.126	29	0.296
Gray	24	0.031	4	0.167
Gray br	29	0.037	5	0.172
Tan	14	0.018	2	0.143
Tan br	17	0.022	3	0.176
Brown	93	0.120	52	0.559
Brown br	8	0.010	3	0.375
Brownish black	21	0.027	6	0.286
Brownish black br	2	0.003	0	0
Reddish black	9	0.012	1	0.111
Reddish black br	3	0.004	0	0
Reddish brown	3	0.004	2	0.667
Reddish brown br	2	0.003	2	1.000
Totals	776		234	

[1]29 animals excluded for no phenotypic description available

[2]These values are a proportion of the total animals

[3]These values are a proportion of the color category

Analysis of Gray

Multiple genome-wide association analyses were conducted. In the initial analysis of both phenotypically gray vs. not gray and the phenotypically gray and reddish tinged vs. not gray, *melanocortin 1 receptor* (*MC1R*) genotypes were not included as covariates. This resulted in a peak in the *P* values at BTA18, where *MC1R* is located and is known to be responsible for base color. After taking into consideration the *MC1R* genotypes and fitting for that effect, the peaked *P* values shifted to BTA6. Figure 2 and Figure 3 show the $-\log_{10}(P)$ plotted for each and after removing any SNP marker that had an undetermined location. A single significant SNP marker was determined in the gray vs. not gray analysis at a location of 68,059,441 bp ($P_{raw} = 9.69 \times 10^{-7}$, $P_{adjusted} = 0.032$) (UMD_3.1, NCBI project 32899, Gen Bank GCA_000003055.3) on BTA6. In the second analysis of gray, which analyzed phenotypically gray and cattle that were light in color but had a reddish tinge vs. not gray, 5 significant markers on BTA6 were identified, forming a region from 62.93 Mb to 83.92 Mb (UMD_3.1, NCBI project 32899, Gen Bank GCA_000003055.3). The same lead SNP marker from the first analysis was present in the second, but had an increased significance ($P_{raw} = 1.50 \times 10^{-10}$, $P_{adjusted} = 5.02 \times 10^{-6}$). These regions on BTA6 ruled out previous candidate genes that cause gray in other species, *premelanosome protein* (*PMEL*), located on BTA5, and *syntaxin-17* (*STX17*), located on BTA8. Genes such as *v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog* (*KIT*), which is known to cause white coloration (spotting), *platelet-derived growth factor receptor alpha polypeptide* (*PDGFRA*), the strongest candidate gene for the reddening phenotype in Nellore-Angus cattle (Hulsman

Hanna et al., 2014), and *corin serine pepsidase* (*CORIN*), lighter color in mice, all lie within the interval created and were not able to be ruled out.

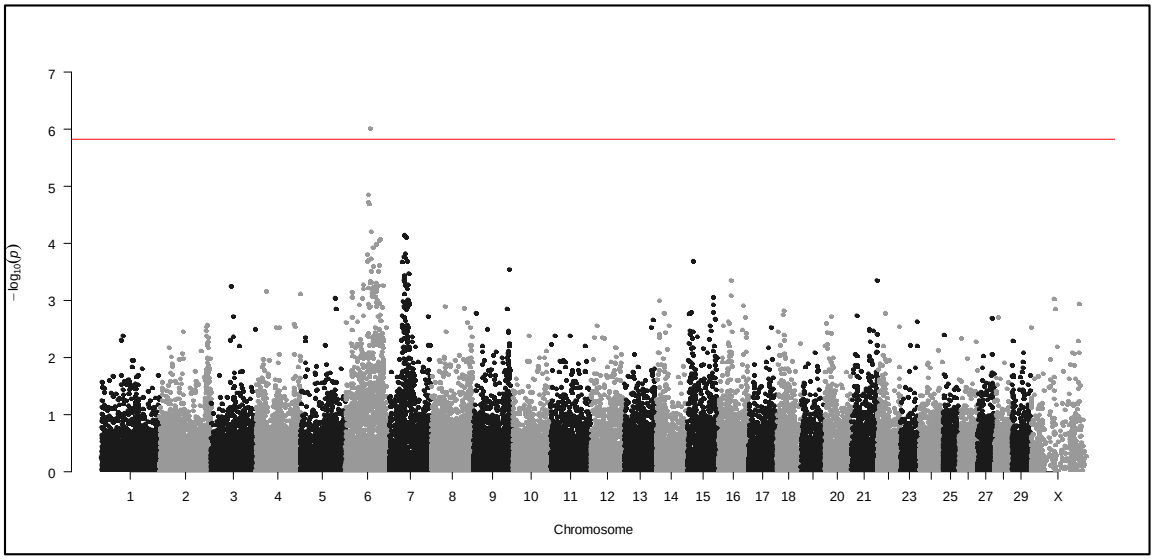

Figure 2. The $-\log_{10}(P)$ of genome-wide association analysis of gray vs. not gray. Fitted for genotypes at *MC1R*. The $-\log_{10}(P)$ is plotted for each SNP. The horizontal red line indicates the cut off for $P = 0.05$ after Bonferroni correction.

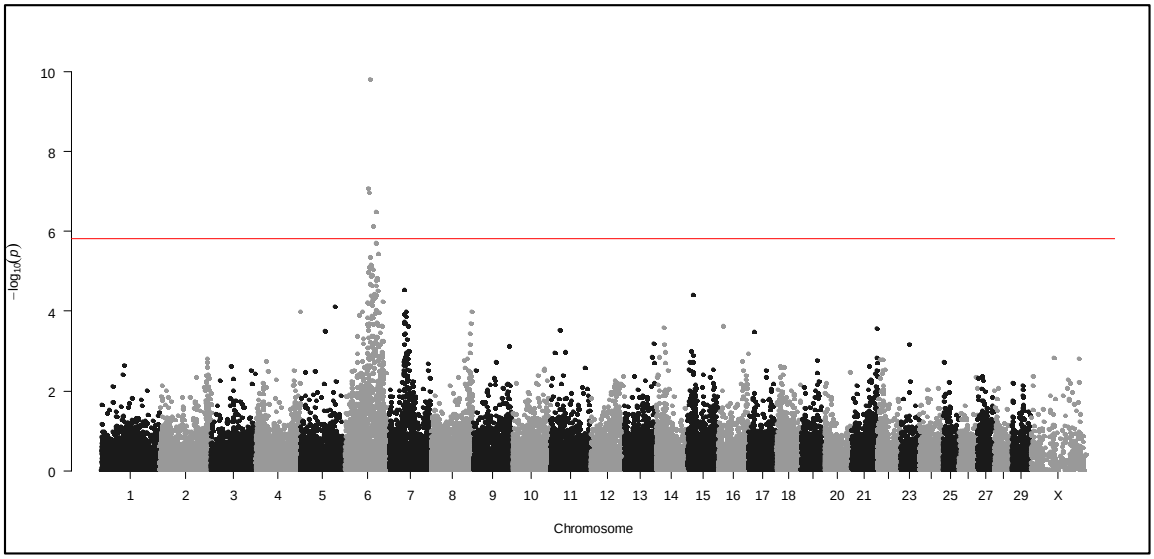

Figure 3. The $-\log_{10}(P)$ of genome-wide association analysis of gray and reddish tinged vs. not gray. Fitted for genotypes at *MC1R*. The $-\log_{10}(P)$ is plotted for each SNP. The horizontal red line indicates the cut off for $P = 0.05$ after Bonferroni correction.

Another analysis was conducted using the residuals of degree of black coloration scores as described by Hulsman Hanna et al. (2014) to try to establish whether reddening and gray are a single locus or two different loci. These values were fit as covariates as well as the genotypes at *MC1R*. The analysis resulted in the estimate of the proportion of variance in phenotypes explained by typed genotypes to be near zero (regression line has no predictive value), therefore the analysis could not be completed, suggesting that gray and reddening are the caused by the same gene or are very closely linked.

SUMMARY

Phenotypic Proportions

Cattle were classified into sixteen different groups, black, red, gray, tan, brown, brownish black, reddish black, reddish brown, and the brindle variants of each, based upon available records. The values calculated were taken as the fraction of that color group out of the population in consideration. The spotting values were taken as the fraction of the number of spotted animals with in that color group out of all animals in that color group. Approximately half of the population had a black phenotype, but based upon observed inheritance, approximately three-fourths of the population should be black.

Analysis of Gray

Using GEMMA (Zhou and Stephens, 2012), genome-wide association analyses were conducted on the cattle. Without fitting for *melanocortin 1 receptor* (*MC1R*) genotypes, *P* values peaked at BTA18, where *MC1R* is located and is responsible for base color determination (red vs. black). When genotypes at *MC1R* were fitted as covariates, *P* values peaked at BTA6 in both the gray vs. not gray analysis and the gray and reddish tinged vs not gray analysis. There was one significant SNP marker in the first analysis and 5 significant markers in the second, creating a large region on BTA6 between 62.93 Mb to 83.92 Mb (UMD_3.1, NCBI project 32899, Gen Bank GCA_000003055.3). The significant SNP markers on BTA6 ruled out previous candidate genes of gray from other species. Fitting the reddening scores from Hulsman

Hanna et al. (2014) as fixed effects along with the *MC1R* genotypes could not be analyzed because of the proportion of variance in phenotypes explained by genotypes was nearly zero, meaning that the regression line has no predictive value.

CONCLUSION

The results of the analyses of gray and the region created on BTA6 ruled out previous gray candidate genes from other species, *syntaxin-17* (*STX17*) in horses, located at BTA8 in cattle, and *premelanosome protein* (*PMEL*) in humans and mice, located at BTA5 in cattle. Both *v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog* (*KIT*), which is known to cause white coloration (spotting), and *platelet-derived growth factor receptor alpha polypeptide* (*PDGFRA*), the strongest candidate gene for the reddening phenotype in Nellore-Angus cattle (Hulsman Hanna et al., 2014) are in the region on BTA6 determined from the analyses. *Corin serine pepsidase* (*CORIN*), known to cause light coat color in mice, also lies within the interval and coincides with the lead SNP marker from both analyses. The proportion of variance in phenotypes expected by genotypes being nearly zero, when residuals of reddening scores were fitted as covariates, suggest that reddened (Hulsman Hanna et al., 2014) and gray phenotypes are both caused by the same gene or the causative genes are very closely linked.

LITERATURE CITED

Albrecht, E., K. Komolka, J. Kuzinski, and S. Maak. 2012. Agouti revisited: Transcript quantification of the ASIP gene in bovine tissues related to protein expression and localization. PLoS One. 7(4): e35282. doi:10.1371/journal.pone.0035282.

Barendse, W., D. Vaiman, S. J. Kemp, Y. Sugimoto, S. M. Armitage, S. M. Williams, H. S. Sun, A. Eggen, M. Agaba, S. A. Aleyasin, M. Band, M. D. Bishop, J. Buitkamp, K. Byrne, F. Collins, L. Cooper, W. Coppietters, B. Denys, R. D. Drinkwater, K. Easterday, C. Elduque, S. Ennis, G. Erhardt, L. Ferretti, N. Flavin, Q. Gao, M. Georges, R. Gurung, B. Harlizius, G. Hawkins, J. Hetzel, T. Hirano, D. Hulme, C. Jorgensen, M. Kessler, B. W. Kirkpatrick, B. Konfortov, S. Kostia, C. Kuhn, J. A. Lenstra, H. Leveziel, H. A. Lewin, B. Leyhe, L. Lil, I. Martin Burriel, R. A. McGraw, J. R. Miller, D. E. Moody, S. S. Moore, S. Nakane, I. J. Nijman, I. Olsaker, D. Pomp, A. Rando, M. Ron, A. Shalom, A. J. Teale, U. Thieven, B. G. D. Urquhart, D-I. Vage, A. Van de Weghe, S. Varvio, R. Velmala, J. Vikki, R. Weikard, C. Woodside, J. E. Womack, M. Zanotti, and P. Zaragoza. 1997. A medium-density genetic linkage map of the bovine genome. Mamm. Genome. 8:21–28.

Berryere, T. G., S. M. Schmutz, R. J. Schimpf, C. M. Cowan, and J. Potter. 2003. *TYRP1* is associated with dun coat colour in Dexter cattle or how now brown cow? Anim. Genet. 34:169–175.

Bland, J. M., and D. G. Altman. 1995. Multiple significance tests: the Bonferroni method. BMJ. 310:170.

Dreger, D. L., and S. M. Schmutz. 2010. The variant red coat colour phenotype of Holstein cattle maps to BTA27. Anim. Genet. 41:109–112.

Drögemüller, C., M. Engensteiner, S. Mose, S. Rieder, and T. Leeb. 2009. Genetic mapping of the belt pattern in Brown Swiss cattle to BTA3. Anim. Genet. 40:225–229.

Drögemüller, C., S. Demmel, M. Engensteiner, S. Rieder, and T. Leeb. 2010. A shared 336 kb haplotype associated with the belt pattern in three divergent cattle breeds. Anim. Genet. 41: 304–307.

Durkin, K., W. Coppieters, C. Drögemüller, N. Ahariz, N. Cambisano, T. Druet, C. Fasquelle, A. Haile, P. Horin, L. Huang, Y. Kamatani, L. Karim, M. Lathrop, S. Moser, K. Oldenbroek, S. Rieder, A. Sartelet, J. Sölkner, H. Stålhammar, D. Zelenika, Z. Zhang, T. Leeb, M. Georges, and C. Charlier. 2012. Serial translocation by means of circular intermediates underlies colour sidedness in cattle. Nature. 482:81–86.

Enshell-Seijffers, D., C. Lindon, and B. A. Morgan. 2008. The serine protease Corin is a novel modifier of the agouti pathway. Development. 135(2):217–225. Epub. 2007.

Girardot, M., S. Guibert, M. P. Laforet, Y. Gallard, H. Larroque, and A. Oulmouden. 2006. The insertion of a full-length *Bos taurus* LINE element is responsible for transcriptional deregulation of the Normande *Agouti* gene. Pigment Cell Res. 19:346–355.

Grosz, M. D., and M. D. MacNeil. 1999. The "spotted" locus maps to bovine chromosome 6 in a Hereford-cross population. J. Hered. 90:233–236.

Hulsman Hanna, L. L., J. O. Sanders, D. G. Riley, C. A. Abbey, and C. A. Gill. 2014. Identification of a major locus interacting with *MC1R* and modifying black coat color in an F_2 Nellore-Angus population. Genet. Select. Evol. 46:4.

Joerg, H., H. R. Fries, E. Meijerink, and G. F. Stranzinger. 1996. Red coat color in Holstein cattle is associated with a deletion in the *MSHR* gene. Mamm. Genome. 7:317–318.

Kavar, T., and E. Čeh, P. Dovč. 2012. A simplified PCR-based method for detection of gray coat color allele in horse. Mol. Cell. Probes. 26:256–258.

Klungland, H., D. I. Våge, L. Gomez-Raya, S. Adalsteinsson, and S. Lien. 1995. The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination. Mamm. Genome. 6:636–639.

Kühn, C., and R. Weikard. 2007. An investigation into the genetic background of coat colour dilution in a Charolais x German Holstein F_2 resource population. Anim. Genet. 38:109–113.

Kwon, B. S., C. Chintamaneni, C. A. Kozak, N. G. Copeland, D. J. Gilbert, N. Jenkins, D. Barton, U. Francke, Y. Kobayashi, and K. K. Kim. 1991. A melanocyte specific gene, Pmel 17, maps near the silver coat color locus on mouse chromosome 10 and is in a syntenic region on human chromosome 12. Proc. Natl. Acad. Sci. USA. 88:9228–9232.

Martínez-Esparza, M., C. Jiménez-Cervantes, D. C. Bennett, J. A. Lozano, F. Solano, and J. C.García-Borrón. 1999. The mouse *silver* locus encodes a single transcript truncated by the *silver* mutation. Mamm. Genome. 10:1168–1171.

Norris, B. J., and V. A. Whan. 2008. A gene duplication affecting expression of the ovine *ASIP* gene is responsible for white and black sheep. Genome Res. 18:1282–1293.

Olson, T. A. 1999. Genetics of colour variation. In: R. Fries and A. Ruvinsky, ed., The genetics of cattle. CAB International, Wallingford, UK. p. 33–53.

Philipp, U., B. Lupp, S. Mömke, V. Stein, A. Tipold, J. C. Eule, J. Rehage, and O. Distl. 2011. A *MITF* mutation associated with a dominant white phenotype and bilateral deafness in German Fleckvieh cattle. PLoS ONE. 6(12): e28857. doi:10.1371/journal.pone.0028857.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81:559–575.

Reinsch, N., H. Thompson, N. Xu, M. Brink, C. Looft, E. Kalm, G. A. Brockmann, S. Grupe, C. Kühn, M. Schwerin, B. Leyhe, S. Hiendleder, G. Erhardt., I. Medjugorac, I. Russ, M. Förster, R. Reents, and G. Averdunk. 1999. A QTL for the degree of spotting in cattle shows synteny with the *KIT* locus on chromosome 6. J. Hered. 90:629–634.

Rhoad, A. O. 1936. The silver gray color in Indian cattle. J. Hered. 27:113–118.

Riley, D. G., T. H. Welsh Jr., C. A. Gill, L. L. Hulsman, A. D. Herring, P. K. Riggs, J. E. Sawyer, and J. O. Sanders. 2013. Whole genome association of SNP with newborn calf cannon bone length. Livest. Sci. 155:186–196.

Schläpfer, J., N. Stahlberger-Saitbekova, J. E. Womack, C. Gaillard, and G. Dolf. 2001. Assignment of six genes to bovine chromosome 13. J. Anim. Breed. Genet. 118:189-196.

Schmidtz, B. H., F. C. Buchanan, Y. Plante, and S. M. Schmutz. 2001. Linkage mapping of the tyrosinase gene to bovine chromosome 29. Anim. Genet. 32:119–120.

Schmutz, S. M., and D. L. Dreger. 2013. Interaction of *MC1R* and *PMEL* alleles on solid coat colors in Highland cattle. Anim. Genet. 44:9–13.

Schmutz, S. M., T. G. Berryere, D. C. Ciobanu, A. J. Mileham, B. H. Schmidtz, and M. Fredholm. 2003. A form of albinism in cattle is caused by a tyrosinase frameshift mutation. Mamm. Genome. 15:62–67.

Seitz, J. J., S. M. Schmutz, T. D. Thue, and F. C. Buchanan. 1999. A missense mutation in the bovine *MGF* gene is associated with the roan phenotype in Belgian Blue and Shorthorn cattle. Mamm. Genome. 10:710–712.

Sponenberg, D. P. 2009. Equine color genetics. 3rd ed. Iowa St. Press, Ames, IA.

Wasmeir, C., A. N. Hume, G. Bolasco, and M. C. Seabra. 2008. Melanosomes at a glance. J. Cell Sci. 121:3995–3999.

Werth, L. A., G. A. Hawkins, A. Eggen, E. Petit, C. Elduque, B. Kreigesmann, and M. D. Bishop. 1996. Rapid communication: melanocyte stimulating hormone receptor (*MC1R*) maps to bovine chromosome 18. J. Anim. Sci. 74:262.

Zhou, X., and M. Stephens. 2012. Genome-wide efficient mixed-model analysis for association studies. Nat. Genet. 44:821–824.